



Text Readability: A Snapshot

Rafizah Mohd Rawian¹

¹School of Languages, Civilisations and Philosophy, Universiti Utara Malaysia, Malaysia

ARTICLE INFO

Article history:

Received Sep 28, 2018

Revised Jan 02, 2018

Accepted Jan 02, 2019

Keywords:

*Text Readability Analysis,
Readability Formula,
Text Difficulty*

Conflict of Interest:

None

Funding:

None

Corresponding Author: Rafizah Mohd Rawian, Senior Lecturer at School of Languages, Civilisation and Philosophy, Universiti Utara Malaysia, Malaysia. Email: rafizah@uum.edu.my Tel. +60123630455

Copyright © Association of Language Teachers in Southeast Asia.
All rights reserved

ABSTRACT

Selecting suitable reading materials are taxing and challenging for many English instructors. Text readability analysis can be used to automate the process of reading material selection and also the assessment of reading ability for language learners. Readability formulas have been broadly used in determining text difficulty based on learners' grade level. Based on mathematical calculations, a readability formula examines certain features of a text in order to provide best rough approximations as an indication of difficulty. This paper reflects some aspects and issues of readability analysis

1. DEFINITION OF READABILITY

The concept of readability was initially defined as the interactions of understanding of words, phrases and ideas in a text with comprehension that influence readers' interest, understanding and reading speed (Dale & Chall, 1949). Readability is further explained as "the ease of understanding or comprehension due to the style of writing" (Klare, 1963). This definition highlights the importance of writing style besides aspects like content, coherence and organization. According to McLaughlin (1969), readability signifies the interaction between the text and a group of readers of known characteristics like reading skill, prior knowledge, and motivation. All in all, readability is an attribute of written text, commonly defined by factors that theoretically make a text more or less difficult to read (Begeny & Greene, 2013).

2. DEVELOPMENT OF READABILITY FORMULAS

Readability formulas were first initiated between 1921 and 1934 where vocabulary was the primary basis to predict readability. Between 1934 and 1953, formulas like Flesch (1948), Dale and Chall (1948), Gunning (1952), Spache (1953) and Powers-Sumner-Kearl (1958) that encompassed more and different factors as variables, with less dependency on the Thorndike word count were developed.

The arrival of cloze procedure as a tool for measuring readability in the mid-1950s incited the development of new criteria, new formulae, computerized versions, and the continued testing of text variables (Dubay, 2004). The cloze testing was considered to be relevant in measuring not only the difficulty of the whole written text but also the difficulty of individual words, phrases, and clauses (Bormouth, 1966). Automated Readability Index - ARI (1967), Bormuth Readability Index (1968), Fry's Readability Formula (1968), McLaughlin's (1969) *Simple Measure of Gobbledygook* (SMOG), FORCAST Readability Formula (1973), Coleman-Liau Readability Formula (1975), Raygor Estimate Graph (1977) and the new Chall and Dale (1995) were later introduced. Since then, more computerized formulas have been developed. Some recent online readability tools

include Lexile Analyzer, Lexile Framework, ATOS, LexTutor, Text Analyzer, Vocab Kitchen, Coh-Metrix and Coh-Metrix Common Core Text Ease and Readability Assessor (T.E.R.A). Microsoft Word also provides inbuilt readability measures - the Flesch Reading Ease (1948) and the Flesch-Kincaid Grade Level Formulas (1975).

Since most readability formulas predict reading difficulty for native readers, several online readability tools have been inaugurated to cater ESL/FL learners and also to analyse foreign texts. McAlpine EFLAW is used to determine the suitability of English texts for ESL learners. Cambridge's Text Inspector (2016) can instantly analyse the difficulty levels of English texts using the Common European Framework of Reference for Languages (CEFR). Its EVP (English Vocabulary Profile) is an online vocabulary resource that contains information about which words and phrases are acquired by learners at each CEFR level (Xia, Kochmar & Briscoe, 2016). LIX is a Swedish readability formula that calculates the difficulty of reading texts in Swedish, French, German and Greek. Spanish Lexile Analyzer is also available to analyse Spanish texts.

3. APPLICATION OF READABILITY FORMULAS

Readability can be assessed through readability tests that apply readability formulas. These formulas are mathematical in nature; their primary aim is to measure the grade level a person must have to read and comprehend a text (Grabe, 2009). Readability measures are based on two features of a text - the difficulty of the sentences (usually measured by the number of words or clauses per sentence) and the familiarity of the words (usually measured by the frequency of the words in a large database of texts). To estimate the difficulty of a text, readability formulas commonly use the length of a word or sentence as the indicator of difficulty. It is assumed that longer sentences are harder sentences and longer words are harder words. The computations would calculate the actual scale score and this score would then be compared and interpreted based on an index table determining readability scores (Madelaine & Wheldall, 2004). It should be noted that each readability formula has its own index table and this table differs from one another. Below are some examples of index tables used by different readability formulas.

Example 1: The new Dale-Chall Readability Formula

Dale-Chall Raw Score to Grade Conversion

Formula Raw Score	Grade Level
4.9 and below	4 th Grade and below
5.0 to 5.9	5-6 th Grade
6.0 to 6.9	7- 8 th Grade
7.0 to 7.9	9-10 th Grade
8.0 to 8.9	11-12 th Grade
9.0 to 9.9	13-15 th Grade (College)
10.0 and above	16 th Grade and above (College Graduate)

Source: Dale, E. & Chall, J.S. (1995). *Readability revisited: The new dale-chall readability formula. Educational Research Bulletin*, 78, 214-223.

Example 2: The Flesch Reading Ease Formula

The Flesch Readability Index

Flesch Index	Educational Level
91-100	5 th Grade
81-90	6 th Grade
71-80	7 th Grade
66-70	8 th Grade
61-66	9 th Grade
51-60	High School
31-50	Some College
0-30	College Graduate
<0	Law School Graduate

Source: Flesch, R.F. (1948). *A new readability yardstick. Journal of Applied Psychology*, 32, 221-233

There are many readability measures and though each is slightly different, all of them focus solely on the surface difficulty of the words and the sentences. Applying different readability tools to the same piece of text will provide different scores and different reading levels (Begeny & Greene, 2013). The results of testing therefore depend significantly on which readability tool is used.

4. SIGNIFICANCE OF READABILITY FORMULAS

Studies have shown that readability formulas are useful and have been widely accepted. Fry (1968), for instance, has stated that a lot of “good deal has been written about them” (p.513). Readability tools are helpful in objectively analysing texts and this veracity has made them more likely to be reliable and trusted. Rubin and Bolt (1981, p.5) reaffirm that “readability formulas are available, objective, economical, and established”.

Print materials used in classrooms may benefit from the use of an objective quantitative measurement of readability. Rush (1984, p.4) states that “readability scores can be used to determine the grade-level appropriateness of materials ranging from library books and periodicals, to instructional materials in subjects such as social studies, science, mathematics, health, and even reading”.

The formulas provide unbiased results that language instructors can trust to predict the level of readability in broad terms. The readability scores would give rough estimations as to whether the selected materials are pitched at a level of readability appropriate for the reading stage of the students. The outcome of the readability analysis not only minimizes time and effort but also avoids the “trial and error” reading practice in deciding the text difficulty.

Since readability formulas are text-based formulas; they are easy-to-use. Calculations using the formulas can be done by hand or by using computerized versions that are built into word processing programs or other software. The online readability tools are very user-friendly. A user is only required to paste a sample or an excerpt of a text (150-3000 words) in a box and clicks on a button to analyse a text. The online readability tools will then analyse the text and output the results based on several readability formulas. These tools will also assist the user to determine the grade level for the text. *Automatic Readability Checker*, *Readable* and *Readability Calculator* are some of the online tools that can assess texts using several renowned formulas such as Flesch-Kincaid Grade Level, the Gunning-Fog Index, Coleman-Liau Index, ARI (Automated Readability Index), and SMOG in a simultaneous manner.

Readability tools can also inform whether any changes that have been made to the text are really making it clearer. In addition, readability formulas can also enhance readers’ retention, comprehension and reading speed.

5. DOWNSIDES OF READABILITY FORMULAS

Despite the progressive and rapid development of the readability formulas, they still remain deficient. One of the shortcomings of these formulas is their incapability to cater struggling readers, learning-disabled readers and also ESL/EFL learners. In addition, there is a tendency that the readability results may underestimate the difficulty of the reading materials (Mesmer, 2008).

Another constraint of the readability formulas is their ignorance towards other factors that may ease reading and influence comprehension. According to Armbruster et al., (1985, p.8) “readability formulas fail to take into account many characteristics of a text that are known to affect comprehension, for instance content difficulty and familiarity, organization of ideas, author style, page layout”. In addition, the formulas also ignore the active role of the reader thus they can only determine the level of readability for a wide range of students in general.

Another drawback is the inconsistent readability scores for the same text. Readability formulas vary in which attributes of words and sentences they take into account and how they measure them. For instance, the Flesch Reading Ease Score calculates the total number of words in each sentence, and then the total number of syllables in each word while the Automated Readability Index (ARI) calculates based on characters per word instead of syllables. With different formulas, the scores for the same text can sometimes differ by two, three, or more grade levels (Pitcher & Zhihui, 2007). For example, Flesch-Kincaid scores tend to underestimate actual reading

grade level because they are often several grade levels below the results obtained by other measures (Root & Stableford, 1998).

6. CONCLUSION

In order to meaningfully use readability formulas, language instructors should understand and realize what these formulas actually measure as well as their limitations. Readability scores can be used as a quick screen for difficulty of words and sentences but not as indicators of comprehension. The scores exhibit only one of many factors that can influence the ease of reading and the usability of reading materials.

REFERENCE

- Armbruster, B., Osborn, J., & Davison, A. (1985). Readability formulas may be dangerous to your textbooks. *Educational Leadership*, 42(7), 18-20.
- Begeny, J. C. & Greene, D. J. (2013). Can readability formulas be used to successfully gauge difficulty of reading materials? *Reading in Schools*, 00(00), 1- 18.
- Chall, J. (1981). *Readability: Conceptions and misconceptions* [electronic version]. New York: McGraw-Hill. Retrieved from <http://www.eric.ed.gov/>
- Crossley, S, Allen, D, McNamara, D. (2011). "Text readability and intuitive simplification: A comparison of readability formulas N1 English". *Reading in a Foreign Language*. 23 (1): 86–101.
- Dale, E. & Chall, J.S. (1995). Readability revisited: The new dale-chall readability formula. *Educational Research Bulletin*, 78, 214-223.
- Flesch, R.F. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32, 221-233
- Fry, E, (1968). A readability formula that saves time. *Journal of Reading*, 11, 513-516, 575-578.
- Grabe, W. (2009). *Reading in second language: Moving from theory to practice*. New York: Cambridge University Press.
- Graesser, A.C., McNamara, D.S.& Louwerse, M.M (2003). *What do readers need to learn in order to process coherence relations in narrative and expository text*. In A.P. Sweet and C.E. Snow (Eds.), *Rethinking reading comprehension* (pp. 82–98). New York: Guilford Publications.
- Giles, T. D., & Still, B. (2005). A syntactic approach to readability. *Journal of Technical Writing & Communication*, 35(1), 47-70.
- Klare, G. (1963). *The measurement of readability*. Ames, IA: Iowa State University
- Mesmer, H. A. E. (2008). *Tools for matching readers to texts: Research-based practices*. New York: The Guilford Press
- Pitcher, B., & Fang, Z. (2007). Can we trust levelled texts? An examination of their reliability and quality from a linguistic perspective. *Literacy*, 41(1), 43-51.
- Root, J. & Stableford, S. (1998). *Write it easy-to-read: A guide to creating plain English materials*. Health Literacy Centre, University of New England, Biddeford, Maine.
- Rubin, A. & Bolt, B. (1981). *Conceptual readability: New ways to look at text* (Reading Education Report No. 31). Urbana, Illinois: Centre for the Study of Reading. ERIC Documentation Reproduction Service Number ED 208 370.
- Rush, R. (1984). Assessing readability: Formulas and alternatives. *Reading Teacher*. 39(3). 274-283.
- Zhihui, F., & Schleppegrell, M. J. (2010). Disciplinary literacies across content areas: Supporting secondary reading through functional language analysis. *Journal of Adolescent & Adult Literacy*, 53, 587-597